UNCLASSIFIED

AD NUMBER
AD871612
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; JUN 1970. Other requests shall be referred to Naval Training Service Center, Orlando, FL.**
AUTHORITY
NTDC ltr 9 Dec 1976

THIS PAGE IS UNCLASSIFIED

AD871612

Technical Report: NAVTRADEVCEN 68-C-0371-1

STUDY OF TRAINING EQUIPMENT AND INDIVIDUAL DIFFERENCES: RESEARCH ON INTERACTIVE RELATIONSHIPS AMONG LEARNER CHARACTERISTICS, TYPES OF LEARNING, INSTRUCTIONAL METHODS, AND SUBJECT MATTER VARIABLES

G. Kasten Tallmadge
James W. Shearer

American Institutes for Research in the Behavioral Sciences
Palo Alto, California

Contract N61339-68-C-0371

June 1970

DoD Distribution Statement
This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Commanding Officer, Naval Training Device Center, Code 423, Orlando, Florida 32813.

NAVAL TRAINING DEVICE CENTER

ORLANDO, FLORIDA

NAVTRADEVCEN 68-C-0271-1

STUDY OF TRAINING EQUIPMENT AND INDIVIDUAL DIFFERENCES: Research on Interactive Relationships Among Learner Characteristics, Types of Learning, Instructional Methods, and Subject Matter Variables

ABSTRACT

This study was the fourth in a series of research projects aimed at determining whether learning might be enhanced by employing instructional methods which differed in design and use as a function of learner characteristics. Based on inferences drawn from studies in this series and other research literature, a model was developed enabling the simultaneous examination of the effects of learner characteristics, types of learning, instructional methods, and subject matter variables on achievement.

Each of six experimental courses was administered to between 57 and 60 Navy enlisted men who were previously tested with instruments which yielded a total of 39 measures of aptitude, interest, and personality characteristics.

Correlation and regression analyses revealed no consistent and meaningful interactive relationships existing between learner characteristics and types of learning or types of subject matter. These analyses did, however, reveal an apparently consistent and meaningful interaction between learner anxiety level and method of instruction (inductive vs. deductive). While the magnitude of the observed relationship was not sufficient to give promise of immediate practical application, it was concluded that the research supported the existence of individual differences in learning style.

Recommendations were made for further research based on theoretical formulations involving noncognitive learner characteristics.

FOREWORD

This report describes the results of a project entitled, "Training Equipment and Individual Differences." The purpose of the project, which is part of a Human Factors Laboratory long-range program of applied research on learning, retention and transfer being conducted on both an in-house and contractual basis, was to determine whether employing training systems which differ in design and use as a function of differences in trainee characteristics results in increased training efficiency.

Three earlier phases of this project were reported in the following Technical Reports by G. Kasten Tallmadge and James W. Shearer:

NAVTRADEVCEN 66-C-0043-1, March 1967 (AD 650850)
NAVTRADEVCEN 67-C-0114-1, May 1968 (AD 671842)
NAVTRADEVCEN 67-C-0114-2, August 1968 (AD 674428)

Copies of these reports can be purchased from either the Defense Documentation Center or the Clearing House for Federal Scientific and Technical Information, using the AD number appearing at the end of the reference.

Gene S. Micheli

GENE S. MICHELI, Ph.D.
Human Factors Laboratory

ACKNOWLEDGEMENTS

This project involved more than the usual number of difficulties and frustrations — both technical and administrative. Without the concern, involvement, and particularly the understanding of the project monitor, Dr. Gene S. Micheli, these problems could have been overwhelming. The project staff thanks him sincerely for "keeping the faith" and for providing the kind of support so essential to an effort of this type.

One of the usual ground rules for conducting research in the field is that it shall be done on a "non-interference basis." Anyone involved in this type of research realizes, however, that non-interference is an empty phrase. While such research may not require that a activity's primary mission be abandoned, there is always disruption of established routines and impositions on the personal convenience of many people. For this reason, the investigators wish to thank those who allowed the disruptions to occur and those who sacrificed their personal convenience to support the project. The following is only a partial listing:

Captain J. H. Stevens, Jr., Commanding Officer, and Commander J. O. Ginn, Executive Officer, Service Schools Command, Naval Training Center, San Diego, for their generous support in authorizing A.I.R. to implement the experimental phase of the project in the Service Schools Command.

Mr. Dale Lovell, Educational Advisor, Service Schools Command, Naval Training Center, San Diego, for his assistance in coordinating establishment of the research effort at the Training Center.

Lieutenant W. E. Clark, Officer in Charge of the Basic Electricity and Electronics School, for his assistance in making initial arrangements for the study.

Ensign D. W. Shafer, Chief Warrant Officer W. E. Myers, Senior Chief H. Chaney, Chief J. Rolfes, Chief J. A. Roeser, and Chief S. M. Salerno, Naval Training Center, San Diego, for their valuable assistance in providing experimental *Ss*, classroom facilities and equipment, and Basic Battery scores for all participating *Ss*.

Master Chief W. J. McCoy, Director, Teletype Maintenance School, San Diego, for providing classroom facilities during early stages of the research.

The authors are also indebted to the following persons who provided assistance with the technical and administrative details associated with preparing the experimental courses:

Mr. Frederick W. H. Rothenberg, Naval Training Device Center, San Diego, for contributing information and helpful suggestions regarding Navy training devices.

Mr. Don Hall and Chief A. J. Cappola, Naval Training Device Center, Alameda, for their assistance regarding the training device used in the Celestial Navigation courses.

Commander J. R. Kurfess, Executive Officer, Naval ROTC Unit, Stanford University, for making valuable suggestions and providing materials used in the development of experimental Celestial Navigation courses.

The two authors, of course, did not do all of the work reflected by this report. Again our listing is incomplete, but special thanks are due to:

Mr. A. O. H. Roberts of A.I.R., for helpful contributions relating to statistical analyses.

Mrs. Beverly Laurits of A.I.R., who served as project secretary, administrative assistant, and data analyst. Her diligence and understanding were greatly appreciated.

Mr. Otto t'Hooft of A.I.R., for his valuable assistance in data reduction and analysis.

Finally, thanks are due to Drs. Richard E. Snow and Glenn H. Bracht for their thoughtful review of the preliminary draft of this report. Their comments have helped to clarify several issues in the Introduction and Discussion sections.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SECTION I

INTRODUCTION

Individualization of instruction is considered by many to be the major thrust of educational innovation today. In its broadest and most complete application, such instruction would enable each student (a) to pursue those educational goals most compatible with his unique combination of aptitude, interest, and personality characteristics, (b) to proceed through instructional sequences leading to these educational goals at his own pace, and (c) to work with instructional materials which match, in terms of techniques and media, his particular learning style.

No existing instructional system has incorporated all three of these features with unqualified success — largely because the state-of-the-art has not kept pace with the ideas and ideals of individualized instruction. Progress has been made, however, and new avenues have been opened up which enable students to develop their individual talents in settings designed to bring everyone to his own highest level of achievement.

There are many differences, of course, between training and education. Perhaps most significant is the difference between the strong job orientation of most training programs and the student related objectives frequently reflected by modern educational practice. Clearly, training must ignore that feature of instructional individualization concerned with the selection of topics to be studied.

Individualization through pacing is certainly feasible in many training situations. Still, except where manpower resources are limited or where the objectives are to upgrade certain elements in the labor market, the aim of training program design is generally to achieve a specified set of objectives in minimum time and at minimum cost. Thus, if differentially paced training is offered, it

must be demonstrably superior to standard instruction on some realistic cost-effectiveness criterion.

The third approach to individualizing instruction, namely that of employing training systems which differ in design and use as a function of known differences in characteristics of the trainees, could be employed as well in military and industrial settings as in standard school classrooms. Relationships between learner characteristics and the techniques and media of instruction, however, have not as yet been clearly established. The study reported here represents an attempt to advance the current state-of-the-art with respect to understanding such relationships by employing a more systematic investigation of apparently relevant variables than has been conducted heretofore.

A. History of the Project

The research reported here represents the most recent of a series of experimental studies conducted by the American Institutes for Research and sponsored by the Naval Training Device Center. The purpose of all of these studies has been to investigate the feasibility of enhancing training effectiveness by employing training systems which differ in design and use as a function of known characteristics of learners. The first study of the series (Tallmadge & Shearer, 1967) contrasted two experimental versions of a 1 week course in Maneuvering Board with the standard version of the course in the Radarman Class A School at Treasure Island, California. It produced no significant interactions between instructional method and learner characteristic variables, although some indirect evidence was produced which led Tallmadge (1968) to speculate that the nature of the material to be learned was critical and needed to be examined as a separate, independent experimental variable.

This hypothesis was the basis for the second study (Tallmadge, Shearer, & Greenberg, 1968) which was designed to examine higher

order interactions among learner characteristics, instructional methods, and "subject matters." Inductive and deductive experimental courses were developed for a logico-mathematical subject matter (Transportation Technique) and a visual form discrimination topic (Aircraft Recognition). For both courses interactions occurred between noncognitive learner characteristics and methods of instruction. It was found, however, that the direction of the relationships which existed between learner characteristics and method of instruction for the Transportation Technique courses were reversed for the Aircraft Recognition courses. In other words, those *Ss* who did well in the inductive version of the Transportation Technique course had the same characteristics as *Ss* who did well in the deductive version of the Aircraft Recognition course. Similarly, *Ss* who performed well under the inductive version of the Aircraft Recognition course had the same characteristics as those who performed well under the deductive version of the Transportation Technique course. The relationships found in this research were not only statistically significant but appeared (then) to be of sufficient magnitude to have practical utility.

The two subject matters differed in a number of ways, and it seemed important to speculate as to which of the differences was responsible for reversing the observed relationships between learner characteristics and instructional methods. The experimenters first hypothesized that the meaningfulness of the subject matter was the critical variable and pointed out that the Transportation Technique topic was governed by "rules" which make sense intuitively and have the same kind of truth as equations in physics. The "rules" governing Aircraft Recognition, on the other hand, were described as arbitrary. They could only be learned by rote and did not make sense since they possessed no inherent logic, or truth.

Later, however, as part of an attempt to build a model relating types of subject matter, types of learning, and types of teaching to achievement differences, Tallmadge and Shearer (1968) noted and described how type of subject matter and type of learning had been confounded in the Tallmadge et al. (1968) study. For the "meaningful rules" subject matter (Transportation Technique), the inductive course had taught concepts and principles in addition to problem solving procedures. The deductive course, on the other hand, was limited to coverage of the problem solving procedures. Thus it was not possible to determine whether the observed interaction for the Transportation Technique topic was between learner characteristics and instructional methods or between learner characteristics and types of learning. This unanswered question, plus others raised by study results presented in the literature, formed the basis for the research reported here. A more specific problem statement is presented later in this report.

B. Status of Learning Style Research

Previous reports in this series have attempted to present reasonably complete literature reviews. Two recent reports, however, (Cronbach & Snow, 1969; Bracht, 1969) have reviewed, critiqued, and summarized published studies in a manner far beyond the scope of this report. It was considered, therefore, more appropriate to summarize the conclusions of these authors, making occasional reference to particularly critical studies, rather than attempting to duplicate their efforts.

It was readily apparent from even a cursory review of the two reports cited above that these authors have adopted a more cautious attitude toward the existence of meaningful interactions between learner characteristic and instructional method variables than was prevalent only a year ago. Bracht's literature review summarizes some 90 studies of which only 5 produced interactions which met his

stringent and rather unusual criteria for meaningfulness. According to Bracht, as well as to many other investigators, an interaction in order to be meaningful must be disordinal -- meaning that the regression lines for different treatment groups must cross within the range of experienced predictor scores. (Bracht talks about "treatment lines" as opposed to regression lines. A treatment line is simply a line connecting the plotted mean achievement scores attained by homogeneously grouped learners under the various treatment conditions.) Bracht claims that not only must the treatment lines cross but that the achievement scores attained by each group of learners must differ significantly from treatment to treatment. In other words, for a study involving the sex of the learner and instructional method, one method must be significantly superior for males and the other method significantly superior for females.

Bracht's position with respect to the latter criterion is beyond refutation when the learner characteristic variable is a true dichotomy such as sex. When the learner characteristic of interest is a continuum, such as a general or specific aptitude, consideration of the range of talent included in the experimental groups is clearly relevant. Nonsignificant differences obtained with a restricted range of talent could very well be both statistically and practically significant for a population embodying a wider range of talent.

Also, where continuously variable learner characteristics are being investigated, the manner in which *Ss* are sorted into high and low groups becomes quite critical. If the median predictor score for the entire experimental sample is used as a cutoff point, the probability of finding a significant difference between mean achievement scores is lowered by an amount proportional to the difference between this group median and the point on the continuum of predictor scores where the two regression (or treatment) lines cross. This general consideration has been alluded to by Cronbach and Snow (1969)

but in a different context. They have stressed the importance of reporting absolute scores used to sort Ss into high and low groups (whether or not these scores were group medians) in order that comparisons could be made between different studies employing the same predictor variables.

A second argument against accepting Bracht's criteria for meaningful interactions applies with equal validity to many other authors and concerns the supposed necessity for regression (or treatment) lines to cross within the experienced range of scores. The objection here is that it is seldom possible to develop two instructional treatments involving nontrivial learning which differ only in terms of the one nominal experimental variable.

As most practitioners of programmed instruction will readily verify, the teaching effectiveness of any instructional material can be greatly enhanced through a systematic process of empirical tryout and revision. Rarely, however, are materials developed for research on possible interactive relationships between learner characteristics and instructional variables submitted to such rigorous refinement. It is quite possible, then, that where regression or treatment lines do not cross, it is erroneous to conclude that one type of treatment is superior for all *Ss*. It may simply be that one particular instructional treatment employed in the study was a poor example of the type of treatment it was designed to represent. This point is covered again in the Discussion section of this report.

There is no intention here to discuss in great detail methodological problems associated with research in the area of learner characteristic by instructional treatment interactions. In discussing the fairly negative conclusions reached by Cronbach and Snow (1969) and Bracht (1969), it is important, however, to point out that they may have rejected as meaningless some results which could be

interpreted differently by employing different, but equally reasonable, sets of assumptions.

In summarizing his review of the literature, Bracht reached but few conclusions, and even these he qualified carefully (as he should have done after finding but 5 meaningful interactions in the 90 studies he reviewed). These tentative conclusions, nevertheless, are worth repeating here, at least for purposes of comparison with those of other investigators. He felt that tight control over the treatment variables was an almost necessary, but not sufficient, condition for obtaining learner characteristic by treatment interactions. While studies using laboratory type learning tasks may be somewhat sterile, more studies employing such learning tasks have produced significant interactions than studies employing learning tasks more typical of the classroom. Whether this difference is attributable to the purity of the laboratory learning task, or to the extent of control the experimenter has over the treatment, or both, cannot be determined at this time. Certainly, however, if one were attempting to design a study to investigate the possible existence of learner characteristic by treatment interactions, he would be well advised to use a homogeneous learning task and to maintain tight control over the treatment administration.

In speculating about those learner characteristics most likely to interact with instructional treatment conditions, Bracht was more optimistic than Cronbach and Snow regarding the interactive potential of personality characteristics. With respect to ability variables the two reviews also differed. Bracht felt that factorially simple variables were more likely than factorially complex variables to interact with instructional treatments. Cronbach and Snow held the opposite view.

Two of the five meaningful interactions reported by Bracht involved personality characteristics. These characteristics were

Need for Affiliation (mistakenly called Need for Achievement by Bracht on pp. 140, 149 of his report), and introversion-extroversion. These findings led Bracht to suggest that the personality domain was deserving of further investigation. Cronbach and Snow (1969) were less enthusiastic although Cronbach's (1966) earlier position was that "the interacting variables may have more to do with personality than with ability [p. 90]." In their 1969 report, these authors decried learning style studies involving personality variables describing them as "disappointing and unencouraging [p. 173]." At the same time, however, they described a number of studies reporting significant personality characteristic interactions. In fact, on p. 159 of their report, they indicated that they had found repeated hints in the literature of an interaction between anxiety and instructional treatment variables. This conclusion was consistent with the findings of the Tallmadge et al. (1968) study.

Although neither Cronbach and Snow nor Bracht offered any generalizations about sex differences, this factor has shown up in a sufficient number of studies reporting significant interactions to be worthy of mention here. While the following list is not intended to be exhaustive, interactions between the sex of the learner and instructional variables were reported by Armstrong (1969), Farley and Manske (1969), Ferney (1969), Klausmeier and Quilling (1968), and Tanner (1968). In addition, the Tallmadge et al. (1968) study apparently found an interaction involving the masculinity-femininity personality dimension.

Despite the existence of some apparent consistency in the research literature, an overview would certainly uncover more negative than positive findings and more inconsistencies than consistencies. If different learning styles exist, they are indeed elusive — a conclusion amply demonstrated by the research of Sunderson (1969). After revising an instructional treatment to simplify it and reduce the time consumed by its administration, he found that aptitude-

treatment interactions obtained with the original instruction were reversed for the revised version. A study by Burton and Goldbeck (1962) offers further evidence supporting the unpleasant possibility that reducing the difficulty of the learning task may reverse aptitude-treatment relationships.

C. Statement of the Problem

Many unanswered questions relevant to the learning style issue existed at the time the present project was initiated. It was felt, however, that the most recently completed study of this series (Tallmadge et al., 1968) had established some meaningful relationships which would stand up under replication and which would have practical significance for Navy training programs even before the remaining questions were answered. For this reason, the present study was undertaken largely as an experimental demonstration project adding just a few innovations to the previously employed model to seek answers to specific questions raised by the earlier research.

The primary research concern of the present study was resolving the question of whether type of teaching (inductive vs. deductive) or type of learning (understanding vs. rote)* was responsible for the observed interaction effects. A second research concern was partial verification of the hypothesis that the "meaningful rules — arbitrary rules" dichotomy was responsible for the observed reversal of learner characteristic by instructional method interactions which occurred as a function of training topic differences. Finally, it was hoped that the study would both verify inferences made about relevant learner characteristic variables and sharpen their measurement.

* Type of learning, as that phrase is used here, is certainly not independent of type of teaching (see Tallmadge and Shearer, 1968, for a detailed discussion of this point). This label was chosen, despite its potential for confusion, because other alternatives appeared to be still more confusing.

Aside from these research considerations and the experimental model changes necessitated by their implementation, the study design was highly similar to that of the Tallmadge et al. (1968) study although it involved different training topics. (Note: A somewhat briefer summary of this research was published by Tallmadge and Shearer in 1969). The next section of this report provides specific details on the experimental design.

SECTION II

METHOD

This study was the fourth of a series of research projects aimed at determining whether learning might be enhanced by employing instructional methods which differed in design and use as a function of learner characteristics. It was a logical outgrowth of these studies but reflected the findings of other published research as well.

The basic experimental design model shown in Figure 1 was first described by Tallmadge and Shearer in 1968. As shown in Figure 1, the design permits independent examination of two different subject matters, two different types of learning, and two different types of teaching, as well as interactions among these variables. A fourth dimension, learner characteristics, is not shown on the model but is, of course, an essential ingredient of any research of this type.

		SUBJECT MATTERS	
		Meaningful Rules	Arbitrary Rules
TYPES OF LEARNING	Understanding	1 Expository/ Deductive	
		2 Inductive/ Discovery	
	"Rote"	3 Expository/ Deductive	5 Expository/ Deductive
		4 Inductive/ Discovery	6 Inductive/ Discovery

Figure 1. Model showing relationships among instructional methods, subject matters, and types of learning.

Because there can be no such thing as "understanding learning" of "arbitrary rules" subject matter, the Figure 1 model does not represent a complete factorial design. It may, however, be regarded as two overlapping complete factorial designs. For the "meaningful rules" subject matter, types of learning and types of teaching constitute one 2x2 complete factorial design. Within the "rote" type of learning, subject matters and types of teaching constitute a similar design.

If one or more of the three independent treatment variables proved not to affect relationships between learner characteristics and achievement, treatment groups could be combined in appropriate ways to provide larger sample sizes and thus, more sensitive tests of interactive relationships between learner characteristics and instructional treatment conditions.

A. Selection of Individual Difference Measures

There were two primary bases for the selection of individual difference measures to be used in the present study. Greatest dependence was placed on the findings of earlier studies in this series, but measures which showed promise in other published learning style research studies were also given careful consideration.

For the most part, emphasis was placed on the selection of noncognitive individual difference measures as previous studies in this series (Tallmadge & Shearer, 1967; Tallmadge et al., 1968) strongly suggested that learning styles are independent of either general or specific aptitude and ability traits. Within the noncognitive domain, the Tallmadge et al. study pointed to the following dimensions as particularly relevant: (1) masculinity-femininity (a term traditionally, but perhaps misleadingly, used by psychologists to describe patterns of interests ranging from scientific-technological to social-aesthetic), (2) introversion-extroversion, and (3) anxiety level. No direct measures of these psychological dimensions were

employed in the cited study. They were inferred from the patterns of individual interest and personality scales which appeared in regression equations generated through use of the UCLA BMD-02R Stepwise Regression computer program. It was decided that all scales which appeared in those regression equations would be employed again in the study described here, and that more direct measures of the inferred dimensions would also be used.

Based on these considerations, carry-over tests from the earlier study were the Kuder Vocational Preference Record and the Gordon Personal Profile. Since scores on the Navy Basic Battery tests (General Classification, Arithmetic, Mechanical, and Clerical) were readily available from existing Navy records, these measures were also included. The search for tests to measure the masculinity-femininity personality dimension produced one or two instruments specifically designed for this purpose. Close examination of them, however, revealed that they were both lengthy and dated. They did not appear appropriate for administration to a group of Navy enlisted men in 1969. It was concluded, however, that several readily available personality tests produced scores at least highly correlated with the masculinity-femininity dimension.

Similar considerations applied to tests of anxiety, and again, it seemed more appropriate to select a wide-range personality test rather than one of the highly specific tests of anxiety currently available. Buros' Sixth Mental Measurements Yearbook (1965) was used to identify available instruments which yielded scores appropriate for the purposes of this study as well as to provide some preliminary evaluative data. Candidate instruments were subsequently reviewed by project personnel and the Sixteen Personality Factor Questionnaire was selected as the instrument most likely to provide useful scores relevent to masculinity-femininity and anxiety, as well as other potentially useful scores. After following a similar

selection process, the Eysenck Personality Inventory was selected as providing what appeared to be the most appropriate measure of introversion-extroversion for the purposes of this study.

The Set Toward Education scale (Capretta, Jones, Siegel, & Siegel, 1963), a forced choice test scored on a continuum characterized on the one hand by predisposition to learn isolated facts and on the other, by a predisposition to learn principles and concepts, was also included in the battery. While research using this instrument had not shown it to interact significantly with instructional treatment conditions, it appeared particularly likely to detect interactions between learner characteristics and the understanding vs. rote learning dimension of the present study.

Two other tests selected from the Kit of Reference Tests for Cognitive Factors (French, Ekstrom, & Price, 1963) were included in the battery. While there was some evidence in the literature that the selected Hidden Figures and Gestalt Completion tests showed promise for learning style research, they were included in the battery as much for the purpose of breaking the monotony of the lengthy interest and personality tests as for any other purpose.

Finally, a specially prepared Biographical Inventory was included in the battery since a study by Snow, Tiffin, and Seibert (1965) provided some indication that certain biographical, or personal history, data are predictive of learning styles.*

---

* Biographical instruments of the type used in this study do not yield total scores and must be analyzed in terms of patterns of response to individual items. This type of analysis was considered to be beyond the scope of the present study. At the time the study was undertaken, however, it seemed probable that a working agreement could be arranged with the Psychology Department of Stanford University whereby a graduate student could conduct the required analyses at no cost to the project. Delays were encountered, but the project is now in progress. Should it produce significant results, they will constitute the basis for a supplementary report.

The final battery of individual difference measures, then, consisted of the following instruments:

- Kuder Vocational Preference Record (Form BB)
- Gordon Personal Profile
- Eysenck Personality Inventory (Form A)
- Sixteen Personality Factor Questionnaire (Form B)
- Set Toward Education Scale
- Gestalt Completion Test
- Hidden Figures Test
- Biographical Inventory

B. Selection of Training Topics

Several criteria were established for selection of the two courses to be developed and taught experimentally. First, because so little was known about the conditions under which learning styles could be observed, there were obvious reasons for selecting training topics which differed as little as possible (consistent with other objectives of the research) from those successfully employed in the earlier Tallmadge et al. (1968) study. Minimum requirements were that one course should cover a "meaningful rules" topic and the other an "arbitrary rules" topic and that each should be as homogeneous as possible with respect to types of skills and knowledges required (e.g., factorially simple). It was also necessary that both subject matters be amenable to both an inductive or discovery instructional approach and an expository-deductive approach.

Another criterion related to the selection of courses was that they should be relevant to Navy training programs. While it was not considered essential that the content taught be identical to segments of existing Navy programs, it was felt that sufficient similarity should be provided to permit the study findings to be generalized to Navy applications.

Four additional criteria were generated as a result of practical considerations: (1) the selected courses had to be amenable to group administration, (2) large and/or expensive equipment requirements had to be avoided, (3) prerequisite skills had to be minimal to avoid extensive pre-training or the need to use highly selected subject populations and (4) the topics had to be "new" to the Ss.

Finally, it was considered desirable for the courses to involve the use of existing Navy training devices.

Preliminary discussions held with cognizant NTDC personnel resulted in the identification of a number of training programs and devices which appeared to meet the requirements described above and might prove to be satisfactory vehicles for the experimental demonstration program. Subsequent to these discussions, project personnel consulted with several subject matter and training experts to delineate more specifically the content and nature of the experimental training courses. Based on these discussions, it was concluded that Celestial Navigation and Aircraft Recognition would satisfy all criteria relevant to the selection of material for experimental training.

The first of these topics, Celestial Navigation, was almost entirely numerical in content with only a very few simple plotting requirements. Arithmetic skills (addition, subtraction, and multiplication) were its only prerequisites, yet it was not likely that any of the experimental Ss would have been exposed to it previously.

The second selected topic, Aircraft Recognition, had been employed in the previous year's study and required only visual discrimination skills. The specific experimental training task was to identify aircraft from aerial photographs. Learning of this topic imposed no requirements for acquisition of prerequisite skills or knowledges. The topic also had the advantage that few, if any, of the experimental Ss would have had prior experience in the area.

The contents of the two topics were distinctly different. They required factorially distinct learner abilities, yet each topic was internally homogeneous, and both were at least partially relevant to Navy career fields.

C. Development of Experimental Courses

Courses to be used in this study were developed by project personnel with consulting assistance from subject matter experts. Since the courses covered fairly basic topics, a minimum of consulting aid was necessary. The courses were designed for normal classroom presentation to reflect standard Navy training practices.

Although training time requirements differed somewhat as a function of the subject matter being taught, the two Aircraft Recognition courses were closely equalized in terms of training time as were the four Celestial Navigation courses. Since more material was taught in the understanding Celestial Navigation courses than in the rote courses, additional practice exercises and drills were incorporated into the rote courses to accomplish the desired time equalization.

1. Celestial Navigation

Four experimental Celestial Navigation courses were developed for use in the study: Inductive-Discovery-Understanding, Expository-Deductive-Understanding, Inductive-Discovery-Rote, and Expository-Deductive-Rote. The courses were designed to teach *Ss* how to set up and solve simple navigation problems using procedures adapted specifically for the study.

The four experimental courses covered the following topics: (1) converting local time to Greenwich Mean Time, (2) determining the location (latitude and longitude) of the Geographical Position (GP) of a star, (3) finding the

distance from a ship's true position to the Geographical Position of a star, (4) determining the altitude difference of a star, (5) plotting assumed positions, bearing lines of stars, and lines of position, and (6) obtaining true position fixes. In addition, Ss were presented with a review of the following topics: coordinate plotting, latitude and longitude, angular measurement, and use of a protractor. The courses included lecture presentation, chalkboard presentations, practice problems (handouts), and review. Navy training device 1X3, a "Celestial Navigation Sphere," was used to help explain a number of important navigational concepts in the two "understanding" courses but was not used in the rote courses.

Because of the complexity of the subject matter, course coverage had to be restricted in the following ways in order to stay within reasonable time limits:

- o The use of the sun, moon, and planets for navigational purposes was not included. Coverage was restricted to stars.
- o The azimuths of the Geographical Positions of the observed stars — normally read out of the nine volume H. O. 214 Publication — were "given" to avoid the necessity of providing all students with copies of this publication.
- o Navigational problems were restricted to the Northern Hemisphere and to longitudes between 0 and 180° West.
- o All celestial observations were made exactly on the hour to avoid difficult interpolation problems.
- o The assumption was made that no movement of the navigating vessel occurred between pairs of observations.

- Corrections of observational data for chronometer error and for atmospheric refraction were omitted.
- Certain terminological simplifications were employed.

The simplifications made in the subject matter were pointed out to the students at appropriate places in the courses to minimize possible negative transfer should any of them receive subsequent celestial navigation training.

In order to illustrate more precisely the specific nature of the four Celestial Navigation courses, excerpts from these courses are presented in the Appendix to this report.

2. Aircraft Recognition

Two experimental courses which were employed during the previous phase of the research were used in the present study with only minor modifications. The courses involved a visual discrimination task, i.e., that of identifying aircraft from aerial photographs. One of the courses was designed to be deductive and the other inductive in order that the two courses would be comparable to the two "rote" Celestial Navigation courses with respect to training methods used.

The instructional materials used in the Aircraft Recognition courses were developed in the following manner. Initially, aerial photographs of a large number and variety of aircraft were obtained. Micro-photographs were then taken and aircraft were selected for inclusion in the two experimental courses which provided for variety in terms of size, configuration, etc. Another criterion for inclusion was the availability of several different photographs of each aircraft to ensure variety from the standpoint of ease of identification.

Two sets of black and white slides were next prepared. The first set consisted of slides of portions of aerial photographs designed to show how aircraft appear in aerial photographs. This set of slides also illustrated some of the common problems encountered by photointerpreters which render aircraft identification difficult. Included were examples which showed poor contrast between aircraft and background, distortion, and partial obscuring of aircraft by clouds, hangar roofs, and camouflage. The second set of slides included top view silhouettes of all aircraft to be taught in the experimental courses. This latter group of slides was selected from the Aircraft Recognition Slide Kit (Navy training device 5QQ4). Modifications were made to these slides to eliminate all except top views.

The deductive course also made use of a few additional slides especially selected to illustrate specific wing, fuselage, and tail shapes as well as other special characteristics.

a. <u>The Deductive Course</u>

The deductive Aircraft Recognition course included instruction on a system designed to help *Ss* identify aircraft. The system made use of an arbitrary set of specific recognition features related to wings, horizontal tail surfaces, fuselages, and engines. One additional feature consisted of what was termed unique characteristics.

Sixteen aircraft were taught using the system of recognition features. A slide showing a top view silhouette of each aircraft was presented along with some general interest information and a detailed description of the aircraft in terms of specific recognition

features. After 4 aircraft had been presented individually, slides of the 4 aircraft were used for review. This in turn was followed by a series of practice exercises which *Ss* completed and which were subsequently discussed. The same process was then repeated for the next 4 aircraft with inclusion of an additional practice exercise which contained all the aircraft covered up to that point. This cycle continued until all 16 aircraft had been taught.

b. The Inductive Course

The second Aircraft Recognition course was designed to be taught in an inductive manner. The same 16 aircraft were taught and the presentation included the identical general interest information about these aircraft. However, *Ss* in the inductive course were not taught any system of recognition features nor were the aircraft described in terms of these recognition features, as was the case in the deductive course. Again, aircraft were taught in groups of 4, and instruction on each group was followed by a practice exercise covering all aircraft taught up to that point.

In order to compensate for the course time lost due to not presenting a system for identifying aircraft using specific recognition features, it was necessary to increase the number of times each aircraft slide was presented and the length of exposure during the training and review sessions. *Ss* were shown slides of aircraft and were essentially left to evolve their own system for discriminating between aircraft. When questions arose concerning identifying aircraft presented in practice exercises, project staff members who acted as monitors did not mention wing shapes, number of engines, etc. They did pose questions designed to encourage *Ss* to devise their own system for

identifying aircraft.

D. Development of Criterion Measures

Criterion tests were developed for the two subject matter areas directly from the statements of behavioral objectives and prior to course development. The same criterion test was used for both the inductive and deductive versions of each course.

1. Celestial Navigation

A basic criterion examination was developed which included all types of problems which *S*s had been taught to solve. Examination items were tried out using naive non-professional A.I.R. personnel to obtain information regarding difficulty levels and the amounts of time required to complete items. The examination was revised on the basis of these tryouts and items were arranged approximately in order of increasint difficulty.

Because the solving of Celestial Navigation problems involved several steps, a test scoring system was developed to provide partial credit for partial problem solutions. This scoring system provided a possible raw score range of scores from 0 to 53.

A 20-item true-false test was also developed covering the concepts and principles underlying the problem solving procedures of Celestial Navigation. The rationale behind this test was to ascertain whether *S*s exposed to the "understanding" treatment had indeed understood and whether *S*s not taught these concepts might have "figured them out." It was felt that this information would provide a valuable supplement to the basic criterion test. Unfortunately, however, the test proved to be far too difficult. Not only did the mean scores on this test fail to discriminate among

the four treatment groups, but none of them was significantly different from chance expectations.

2. Aircraft Recognition

The same 56-item criterion examination was employed for the Aircraft Recognition course which had been developed during the previous study of this series (Tallmadge et al., 1968). Each item consisted of an aerial photograph which contained one or more aircraft. The specific aircraft that *Ss* were to identify was circled in black. Four of the aircraft covered in the training were not included in the examination, and five aircraft were included in the examination which were not covered in the course. This procedure was adopted because there was some interest in assessing the effectiveness of different training approaches for the recognition of "new" aircraft. *Ss* were instructed to write the designation of the aircraft in the space provided they knew it and to write an "X" in the blank if they were sure they had not been taught that particular aircraft. The test was scored simply by counting the number of correct responses.

E. Study Implementation

1. Collection of Individual Difference Measures

Ss for the study were 353 Navy enlisted men awaiting assignment to Basic Electricity and Electronics School in San Diego. All the psychological tests described earlier in this report were administered to groups of approximately 30 *Ss* during the first day of the 2-day experimental period. All test administration was conducted by members of the project staff to assure consistency of procedures from testing session to testing session. Two project staff members

were present at all testing sessions. Each testing session lasted approximately 7 hours including "breaks." Scores on the four Basic Battery tests were provided by Naval Training School personnel for each *S*.

2. Course Administration

All experimental courses were presented at Naval Training School facilities with project personnel serving as instructors. Each course was taught to groups of approximately 30 *S*s during the second day of the 2-day testing/training period. Two instructors were present during each administration. The experimental classroom sessions consumed approximately 1 full day (the Celestial Navigation courses required slightly longer than the Aircraft Recognition courses).

3. Collection of Criterion Data

Criterion tests were administered to all *S*s by project staff members immediately following completion of instruction. Again, two instructors were present to facilitate testing and monitoring.

The Celestial Navigation criterion examination required 1 hour and 20 minutes, while the Aircraft Recognition examination required approximately 30 minutes.

SECTION III

RESULTS

A. Correlation and Regression Analyses

Since this study was concerned with the practical as well as the statistical significance of differences in achievement resulting from different treatment and learner characteristic variables, it was intended that mean achievement scores under various experimental conditions would constitute the ultimate basis for evaluating the research results. Because so many variables were involved, however, it was felt that preliminary correlation and regression analyses would serve the very useful purpose of providing guidance for focusing subsequent analyses on a more limited number of experimental variables which showed promise of being relevant. For this reason, correlations were calculated, using the BMD-03D computer program, between predictor and criterion test scores for all 39 predictors in each of the 6 treatment conditions. These correlations are reproduced in Table 1.

Examination of the correlations in Table 1 did not reveal any consistent pattern. With but few exceptions the correlations were of approximately the anticipated magnitude. One notable exception was the correlation ($r = .60$) between the Set Toward Education test and achievement under the Celestial Navigation Inductive Rote treatment condition. The $p = .95$ confidence interval for this correlation was found to be $.41 < p < .74$ indicating a substantial relationship. The corresponding correlations under other treatment conditions were, for the most part, not significantly different from zero.

The Kuder Computational and Scientific scales also showed surprisingly high correlations with achievement ($r = .50$ and $r = .52$, respectively) for the Celestial Navigation Deductive Understanding treatment condition. Neither of these coefficients was significantly different from the corresponding correlations for the other three

TABLE 1

Correlations between Individual Difference Measures and Criterion Test Scores within Each Treatment Group

	Celestial Navigation				Aircraft Recognition	
	Understanding		Rate			
	Induct	Deduct	Induct	Deduct	Induct	Deduct
NAVY GENERAL CLASSIFICATION	.58	.64	.40	.48	.17	.20
NAVY ARITHMETIC	.52	.61	.62	.55	.08	-.05
NAVY MECHANICAL	.12	.07	.30	.17	.29	.07
NAVY CLERICAL	.18	.18	.18	.22	.15	-.06
EYSENCK - EXTRAVERSION	-.17	-.04	-.05	-.20	-.36	-.10
EYSENCK - NEUROTICISM	.03	-.14	.21	-.24	.08	-.11
EYSENCK - LIE SCALE	-.09	-.13	-.06	.01	.06	.09
GESTALT COMPLETION	-.06	.24	.04	-.02	.18	.16
GORDON - ASCENDANCY	.11	-.19	.03	.03	-.09	-.12
GORDON - RESPONSIBILITY	.08	-.03	-.15	.02	-.03	.21
GORDON - EMOTIONAL STABILITY	.06	.02	-.12	.10	.02	.21
GORDON - SOCIABILITY	.00	-.09	-.04	-.08	-.23	-.19
HIDDEN FIGURES	.29	.26	.12	.21	.24	.28
KUDER - MECHANICAL	-.06	-.04	.00	-.04	.23	.22
KUDER - COMPUTATIONAL	.38	.50	.34	.28	-.11	.01
KUDER - SCIENTIFIC	.24	.52	.23	.18	.08	.18
KUDER - PERSUASIVE	-.01	-.07	-.15	-.14	-.18	-.05
KUDER - ARTISTIC	.16	-.17	.14	-.14	-.06	.17
KUDER - LITERARY	.00	-.08	.06	.02	-.03	.01
KUDER - MUSICAL	-.13	-.16	.20	.02	.05	-.17
KUDER - SOCIAL SERVICE	-.18	-.24	-.11	-.13	-.03	-.28
KUDER - CLERICAL	.06	.16	-.14	-.15	-.16	.05
16 PF - OUTGOING	-.08	-.07	.02	-.04	-.29	-.27
16 PF - MORE INTELLIGENT	.15	.15	.11	.34	.11	.20
16 PF - EMOTIONALLY STABLE	-.01	.08	-.26	.21	-.11	.19
16 PF - ASSERTIVE	-.10	.11	.06	.04	.00	.13
16 PF - HAPPY-GO-LUCKY	.06	.06	.10	-.02	-.21	-.04
16 PF - CONSCIENTIOUS	.08	-.04	-.21	-.03	.16	-.10
16 PF - VENTURESOME	-.03	-.14	-.18	-.04	-.29	-.19
16 PF - TENDER-MINDED	.08	-.23	.03	-.08	.16	.08
16 PF - SUSPICIOUS	-.10	.07	.06	-.10	.18	.08
16 PF - IMAGINATIVE	.11	-.04	.20	.00	.25	-.07
16 PF - SHREWD	-.17	-.01	-.03	-.26	-.01	-.26
16 PF - APPREHENSIVE	-.04	.13	.13	-.20	.30	-.20
16 PF - EXPERIMENTING	.06	.25	.36	.26	.15	-.14
16 PF - SELF-SUFFICIENT	.13	.06	.13	.18	.06	.08
16 PF - CONTROLLED	.10	-.02	-.10	-.11	.06	.10
16 PF - TENSE	.06	.00	-.06	-.03	.01	-.12
SET TOWARD EDUCATION	.26	.08	.60	.00	.15	.26

Celestial Navigation treatments, however.

Of primary interest were differences between correlations existing as a function of treatment conditions. No consistent patterns of this type were apparent through visual inspection. Despite these appearances, tests of the significance of differences between corresponding predictor-criterion correlations were made for the following pairs of treatment conditions:

- Celestial Navigation Inductive Understanding vs. Celestial Navigation Deductive Understanding.
- Celestial Navigation Inductive Understanding vs. Celestial Navigation Inductive Rote.
- Celestial Navigation Inductive Understanding vs. Celestial Navigation Deductive Rote.
- Celestial Navigation Deductive Understanding vs. Celestial Navigation Inductive Rote.
- Celestial Navigation Deductive Understanding vs. Celestial Navigation Deductive Rote.
- Celestial Navigation Inductive Rote vs. Celestial Navigation Deductive Rote.
- Celestial Navigation Inductive Rote vs. Aircraft Recognition Inductive.
- Celestial Navigation Inductive Rote vs. Aircraft Recognition Deductive.
- Celestial Navigation Deductive Rote vs. Aircraft Recognition Inductive.
- Celestial Navigation Deductive Rote vs. Aircraft Recognition Deductive.
- Aircraft Recognition Inductive vs. Aircraft Recognition Deductive.

Since there were 39 predictor variables, a total of 429 such tests were made. On the basis of chance expectation alone, between

21 and 22 of these comparisons should have shown "statistically significant differences" at the $p \leq .05$ level. In fact, 26 pairs of correlations were found to be significantly different at this level. Similarly, 4 pairs of correlations which differed at the $p \leq .01$ level were expected on a chance basis while 7 were found.

Several of the paired correlational differences reflected the generally high correlations existing between cognitive predictors and the Celestial Navigation final examination, as contrasted with the low correlations of these predictors with Aircraft Recognition criterion performance. If Navy Basic Battery General Classification and Arithmetic test correlations were excluded from consideration, the number of obtained "statistically significant differences" was below chance expectation. At this level of analysis, then, there appeared to be no meaningful results.

In deference to those authors who have argued that correlational data cannot be meaningfully interpreted and insist that regression coefficients be used in their stead (e.g., Cronbach & Snow, 1969), regression coefficients were calculated corresponding to all the correlation coefficients shown in Table 1. These regression coefficients are presented in Table 2. Differences between pairs of regression coefficients were tested by means of t tests. These tests were performed for all pairs of predictor-criterion relationships for which correlational differences had been tested. The tests of the differences between regression coefficients yielded exactly the same results as the tests of differences between correlation coefficients. The identical pairs of predictor-criterion relationships were found to be significantly different at the same probability levels. The regression analyses added nothing to information obtained from the correlation analyses.

TABLE 2

Regression Coefficients of Criterion Test Scores
on Individual Difference Measures

	Celestial Navigation				Aircraft Recognition	
	Understanding		Rate			
	Index 1	Index 2	Index 1	Index 2	Index 1	Index 2
NAVY GENERAL CLASSIFICATION	.04	1.03	1.09	.63	.30	.29
NAVY ARITHMETIC	.76	.00	1.06	.07	.12	-.04
NAVY MECHANICAL	.16	.09	.06	.19	.30	.09
NAVY CLERICAL	.19	.22	.19	.23	.15	-.09
EYSENCK - EXTRAVERSION	-.30	-.00	-.13	-.50	-.03	-.24
EYSENCK - NEUROTICISM	.08	-.32	.06	-.43	.15	-.28
EYSENCK - LIE SCALE	-.43	-.06	-.32	.05	.30	.50
GESTALT COMPLETION	-.16	.77	.13	-.07	.71	.50
GORDON - ASCENDENCY	.17	-.31	.04	.06	-.14	-.22
GORDON - RESPONSIBILITY	.14	-.06	-.26	.04	-.08	.44
GORDON - EMOTIONAL STABILITY	.06	.03	-.19	.12	.03	.42
GORDON - SOCIABILITY	.08	-.14	-.06	-.11	-.30	-.20
HIDDEN FIGURES	.11	.12	.06	.08	.10	.10
KUDER - MECHANICAL	-.06	-.02	-.00	-.02	.12	.13
KUDER - COMPUTATIONAL	.22	.46	.36	.26	-.10	.01
KUDER - SCIENTIFIC	.15	.33	.16	.09	.06	.13
KUDER - PERSUASIVE	-.01	-.04	-.08	-.08	-.12	-.08
KUDER - ARTISTIC	.09	-.12	.11	-.00	-.08	.10
KUDER - LITERARY	-.00	-.06	.06	.02	-.02	.01
KUDER - MUSICAL	-.12	-.15	.23	.01	.06	-.16
KUDER - SOCIAL SERVICE	-.10	-.16	-.07	-.07	-.01	-.16
KUDER - CLERICAL	.04	.14	-.12	-.12	-.13	.04
16 PF - OUTGOING	-.24	-.19	.06	-.10	-.08	-.72
16 PF - MORE INTELLIGENT	.03	.97	.71	1.08	.56	1.00
16 PF - EMOTIONALLY STABLE	-.03	.23	-.63	.04	-.24	.46
16 PF - ASSERTIVE	-.32	.24	.15	.08	.01	.42
16 PF - HAPPY-GO-LUCKY	.10	.12	.20	-.04	-.46	-.10
16 PF - CONSCIENTIOUS	.32	-.11	-.60	-.00	.46	-.33
16 PF - VENTURESOME	-.04	-.23	-.29	-.06	-.53	-.38
16 PF - TENDER-MINDED	.22	-.76	.09	-.70	.43	.22
16 PF - SUSPICIOUS	-.30	.27	.17	-.25	.56	.21
16 PF - IMAGINATIVE	.27	-.10	.57	.00	.67	-.10
16 PF - SHREWD	-.40	-.03	-.10	-1.06	-.03	-.96
16 PF - APPREHENSIVE	-.08	.19	.20	-.39	.40	-.50
16 PF - EXPERIMENTING	.10	.93	1.10	.96	.52	-.53
16 PF - SELF-SUFFICIENT	.30	.30	.53	.60	.19	.27
16 PF - CONTROLLED	.28	-.07	-.33	-.31	.18	.33
16 PF - TENSE	.12	-.01	-.11	-.06	.03	-.30
YRS FORMAL EDUCATION	.22	.07	.52	.06	.12	.10

H. Replication Analyses

The experimental Aircraft Recognition courses used in this research were essentially identical to the Aircraft Recognition courses used in an earlier study in this series (Tallmadge et al., 1968). The Aircraft Recognition portion of this study, then, constituted one of few replications of a study in which significant learner characteristic by treatment interactions had been found. The earlier study found that the inductive and deductive treatments produced significantly different correlations with the criterion for three Kuder Vocational Interest scales (Computational, Scientific, and Social Service), and two Gordon Personal Profile scales (Ascendancy and Sociability). Not one of the differences between these same pairs of correlations was statistically significant in the present study although, with the exception of Gordon Ascendancy, the direction of the differences was the same in both studies.

In the earlier study, a composite measure consisting of the Kuder Scientific Interest scale (positively weighted), the Gordon Ascendancy scale (negatively weighted), and the Kuder Musical Interest scale (negatively weighted) was developed which maximized the differences among the four treatment groups included in that study. *S*s scoring above the median on this composite measure showed significantly greater (Aircraft Recognition) achievement under the deductive method of instruction than under the inductive method ($p < .001$). *S*s scoring below the median performed better under the inductive treatment condition, although this difference was not statistically significant.

The same composite score was calculated for *S*s taught Aircraft Recognition in the present study. Using these data, no achievement differences between inductive and deductive treatment conditions were found either for *S*s scoring above the median, or those scoring

below. Table 3 presents a comparison of mean achievement scores from the earlier and the present study.

Table 3

Mean Achievement Scores

	Aircraft Recognition Earlier Study		Aircraft Recognition Present Study	
	Inductive	Deductive	Inductive	Deductive
High Group	46.68	53.61	50.16	50.34
Low Group	52.19	48.04	49.17	49.15

No explanation could be found as to why the two replications produced different results. Predictor and criterion test score means and variances were found not to differ significantly between replications nor were the results of the present study (as shown in Table 3) significantly affected by using the composite score median from the earlier study to sort present study *Ss* into high and low groups.

C. Additional Treatment by Learner Characteristic Analyses

As discussed in the Introduction, this study was designed to investigate the effects of three separate experimental variables — subject matters, types of learning, and instructional methods. The correlation and regression analyses described above considered these three variables simultaneously by dealing with predictor-criterion relationships on a within-treatment-group basis. By pooling data from two or more treatment groups, it was also possible to investi-

gate each of the variables separately. The possible effects of type of learning, for example, could be investigated by pooling data from the two Celestial Navigation Understanding courses and comparing the pooled results with results obtained by pooling data from the two Celestial Navigation Rote courses. Analyses of this type were carried out and are described below.

1. Learner Characteristics by Subject Matters

The effect of subject matters was examined by pooling *Ss* from the two Celestial Navigation Rote courses, computing correlations between individual difference measures and criterion test scores for the combined group, and comparing these correlations with similar correlations computed for the combined group of *Ss* from the two Aircraft Recognition courses. Again, tests of the significance of the differences between pairs of correlations were made. Of the 39 pairs of correlations which were examined, 2 were found to be significantly different at the $p < .01$ level, and 3 more at the $p < .05$ level. The 2 larger differences involved cognitive measures — Navy Basic Battery General Classification and Arithmetic tests. These measures showed high positive correlations with the Celestial Navigation courses, and much lower correlations with the Aircraft Recognition courses. This finding was expected and was consistent with similar differences reported by Tallmadge et al. (1968) involving these same two predictors and achievement in courses on the Transportation Technique and Aircraft Recognition.

Noncognitive measures were involved in the other three statistically significant correlational differences. The Kuder Computational scale produced the largest difference

between correlations, showing a significant positive relationship with achievement in the Celestial Navigation courses and a small negative correlation with achievement in Aircraft Recognition. Since the Celestial Navigation courses did involve a significant amount of computation, this finding was also consistent with expectations. The 16 PF Outgoing and Experimenting scales produced the other two significant differences. These relationships were such that "Outgoingness" was negatively related to success in Aircraft Recognition, but unrelated to achievement in Celestial Navigation; while "Experimentingness" was associated with success in Celestial Navigation, but unrelated to achievement in Aircraft Recognition. These relationships did not appear to be particularly meaningful, and no attempt was made to interpret them.

2. Learner Characteristics by Types of Learning

To investigate possible interactions between learner characteristics and types of learning, *Ss* in the two Celestial Navigation Understanding courses were pooled, as were *Ss* in the two Celestial Navigation Rote courses. Correlations were then computed between each of the 39 individual difference measures and criterion test scores for the two composite groups. Differences between the resulting pairs of correlations were then tested. Only 1 of the 39 measures produced a correlation for the understanding learning condition which differed significantly from that produced for the rote learning condition. This measure was the Kuder Musical Interest scale which correlated -.14 and +.12 with achievement for the two types of learning, respectively. Since more than 1

"statistically significant difference" ($p \leq .05$) would have been expected on the basis of chance alone, it had to be assumed that this 1 relationship was not meaningful and, therefore, that no learner characteristics by types of learning interaction existed for the Celestial Navigation subject matter.

3. Learner Characteristics by Instructional Methods

Although there were three inductive and three deductive courses, the instructional methods variable was not initially examined by pooling the *Ss* for the three inductive and the three deductive treatment groups. To avoid possible contamination caused by higher order interactions involving either type of learning or type of subject matter, the decision was made to pool only two groups at a time. Thus the two inductive Celestial Navigation treatment groups were combined, as were the two deductive groups. Similarly, the inductive Celestial Navigation Rote group was pooled with the inductive Aircraft Recognition group, and the deductive Celestial Navigation Rote group was pooled with the deductive Aircraft Recognition group. As with the other analyses, correlations for each composite group were computed between the 39 predictors and criterion test scores. Differences between pairs of these correlations were subsequently tested. Of the 78 comparisons made, 9 statistically significant ($p < .05$) differences were found. The pairs of correlations which were found to be significantly different are presented in Table 4.

TABLE 4

Significantly Different Pairs of Correlations

	Cel. Nav. Understanding plus Cel. Nav. Rote		Cel. Nav. Rote plus Aircraft Recognition	
	Induct	Deduct	Induct	Deduct
Eysenck - Neuroticism	.13	-.19	.14	-.18
Kuder - Artistic	.15	-.13	-	-
16 PF - Emotionally Stable	-.14	.14	-.18	.19
16 PF - Imaginative	-	-	.23	-.03
16 PF - Apprehensive	-	-	.18	-.20
Set Toward Education	.45	.08	.37	.08

Considered individually, the pairs of correlations shown in Table 4 are not particularly meaningful and might, one could assume, be due to chance variations. Three of the scales, however, (Eysenck Neuroticism, 16 PF Emotionally Stable, and 16 PF Apprehensive) are essentially measures of adjustment. The fact that all of these scales are consistent with respect to their interaction with instructional method lends considerable credibility to the existence of an interaction between student adjustment and method of instruction. It might even be argued that the Kuder Artistic scale and the 16 PF Imaginative scale also produce scores which are indicative of emotional adjustment, and since the pattern of correlations of these scales is consistent with those of the other scales, thus offer support for the existence of such an interaction.

The significantly different correlations involving the Set Toward Education test all result from the single very high correlation between this measure and achievement in the Celestial Navigation Inductive Rote treatment group which was discussed above. They neither supported nor detracted from the possible existence of an interaction between method of instruction and the adjustment of learners.

There was no indication that either subject matter or type of learning affected the relationship between instructional methods and learner adjustment. It was decided, therefore, to pool the three inductive treatment groups and the three deductive treatment groups in order to examine the overall effect of instructional method. When this was done, correlations were computed for the three scales clearly related to emotional adjustment. Table 5 presents these correlations as well as multiple correlations for the inductive and deductive groups which were computed by assigning weights of +1 to the Eysenck Neuroticism and 16 PF Apprehensive scales, and a weight of -1 to the 16 PF Emotionally Stable scale.

TABLE 5

Correlations Showing Interactions between Measures of Adjustment and Instructional Methods

	Inductive	Deductive
Eysenck - Neuroticism	.11	-.17
16 PF - Emotionally Stable	-.13	.16
16 PF - Apprehensive	.12	-.16
Multiple	.14	-.20

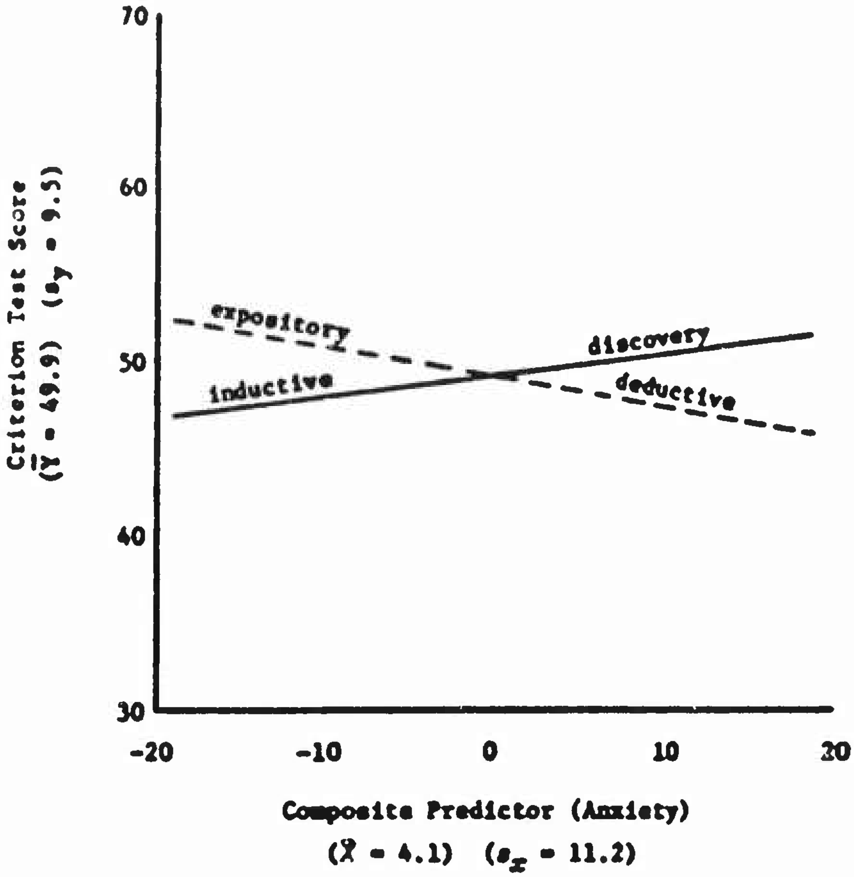

Figure 2. Interaction between anxiety and instructional method.

The regression lines corresponding to the multiple r's of Table 5 are plotted in Figure 2. This figure illustrates that the regression lines cross near the predictor score mean. The slopes of the regression lines were found to be significantly different ($p < .05$) as were the multiple r's. To further investigate this interactive relationship, *Ss*

were subsequently sorted into high and low groups depending upon whether their composite predictor score fell above or below the point where the regression lines crossed (Y = -.36). Finally, mean achievement test scores were computed for those above the crossover point who received an inductive treatment, those above the crossover point who received a deductive treatment, those below the crossover point who received an inductive treatment, and those below the crossover point who received a deductive treatment. These mean scores (together with the associated standard deviations) are presented in Table 6.

TABLE 6

Criterion Test Score Means and Standard Deviations for High and Low Anxiety Groups Taught Inductively and Deductively

	Inductive	Deductive
High Anxiety	$\bar{X}$ = 50.64 s = 9.85	$\bar{X}$ = 48.63 s = 8.94
Low Anxiety	$\bar{X}$ = 48.27 s = 8.91	$\bar{X}$ = 52.14 s = 10.19

Finally, an analysis of variance was performed on the criterion test scores for the four groups using the unweighted means technique (Winer, 1962). The results of this analysis are presented in Table 7. As can be seen, neither

the learner characteristic nor the instructional method main effect was statistically significant while their interaction was significant.

TABLE 7

Unweighted Means Analysis of Variance

Source	*df*	*MS*	*F*	*p*
Instructional Methods	1	69.5	< 1	
Learner Characteristics	1	27.0	< 1	
Interaction	1	710.0	8.0	<.01
Error	349	88.7		

Differences between pairs of means shown in Table 6 were tested using Student's t. The deductive instructional method was found to be significantly superior ($p < .05$) for the low anxiety group. The instructional treatments did not produce significant achievement differences ($p > .10$) for the high anxiety group. Similarly, the low anxiety group performed significantly better than the high anxiety group ($p < .05$) under the deductive treatment while the difference between groups was not statistically significant ($p > .10$) under inductive instructional conditions.

D. Analysis of Overall Treatment Effects

While the primary concern of this study was with interactive relationships involving learner characteristic variables, it did not seem appropriate to ignore completely any overall effectiveness differences which might have existed between treatment conditions. For this reason, mean achievement scores and associated standard deviations

were calculated and are presented in Table 8.

TABLE 8

Mean Achievement Scores and Standard Deviations
for the Six Treatment Conditions

	Inductive	Deductive
Celestial Navigation Understanding	$\bar{X}$ = 50.6 s = 8.8	$\bar{X}$ = 49.0 s = 9.9
Celestial Navigation Rote	$\bar{X}$ = 49.2 s = 10.1	$\bar{X}$ = 51.1 s = 8.8
Aircraft Recognition	$\bar{X}$ = 49.6 s = 9.7	$\bar{X}$ = 49.6 s = 9.7

The F_{max} test was used to test for homogeneity of variance. No significant difference was found (F_{max} = 1.21). Differences between all pairs of means were tested using t tests (a conservative procedure when attempting to minimize Type 1 error). No statistically significant differences were found.

While it should be pointed out that between-subject-matter differences were statistically eliminated by standardizing scores within subject matters, results of the above described tests led to the conclusion that within-subject-matter treatment differences produced no significant achievement differences.

SECTION IV

DISCUSSION

The purpose of this study was to demonstrate the practical significance of learner characteristic by instructional treatment interactions by verifying the experimental findings reported in a previous study (Tallmadge et al., 1968) and by attempting to increase the magnitude of the observed relationships through an improved experimental design and sharpened measurement techniques. It is readily apparent from the results described in the previous section of this report that this goal was excessively optimistic and that far more basic knowledge is required before learning style research is likely to have a significant impact on any real-world training programs.

At the time the study described here was undertaken, other investigators shared with the authors of this report a general feeling of enthusiasm for the practical implications of research in this area. Numerous studies were appearing which reported significant learner characteristic by treatment interactions, and it seemed that soon some order would emerge from the apparently unrelated and frequently discrepant findings. Hypotheses were being formulated and tested in a manner which held promise for establishing a replicable basis for future, more sophisticated refinements.

As the proliferation of research results produced an increasing number of closely comparable studies, the expected emergence of consistent patterns of relationships simply did not occur. Rather, it appeared that relatively minor differences between treatments could cause a complete reversal of interactive relationships. Such reversals were observed by Bunderson (1969) and by Burton and Goldbeck (1962) as a function of differences in the difficulty of learning tasks while subject matter differences produced a similar reversal in the Tallmadge et al. (1968) study (although this difference be-

tween subject matters could hardly be described as minor).

These findings, coupled with the large number of studies reporting negative results or results inconsistent with expectations, have caused the prevailing mood to change from optimism to pessimism. Bunderson (1969), for example, has decried the "robustness" of aptitude-treatment interactions suggesting that other approaches are more likely to improve instructional effectiveness. Certainly the results of the present study are, for the most part, consistent with this recently expressed viewpoint.

Still, it does not appear that learning style research should be abandoned. The relationship found in the present study between anxiety and (inductive vs. deductive) teaching methods, while not impressively large, was sufficiently robust to stand up for two different types of learning, two different subject matters, and probably at least two difficulty levels.

A. <u>Findings of the Present Study</u>

The single positive finding of the present study which the authors feel is reliable and which should stand up to replication was the above mentioned interactive relationship between learner anxiety and method of instruction. While many "statistically significant differences" were found between pairs of predictor-criterion correlation and regression coefficients as a function of treatment conditions, the majority of these differences were considered to be unreliable since they represented only a small fraction of the total number of pairs tested, and since they formed no clear-cut pattern.

The only major exception to the apparently random distribution of predictor-criterion correlation and regression coefficients involved measures of learner adjustment (or conversely, anxiety). Three of the four such scales included in the study (the Gordon Emotional Stability scale was the exception) showed consistent

patterns of interaction with the instructional method variable for both types of learning and subject matters investigated. While but five of the six relationships were statistically significant (based on correlational differences — see Table 4), they were perfectly consistent and indicated that the more anxious students performed better under inductive instructional conditions while less anxious students showed better achievement following deductive instruction. When the three scales were pooled to yield a single anxiety measure, the interaction *F* ratio was statistically significant at the $p < .01$ level, but mean achievement scores were significantly different ($p < .05$) only for *S*s low in anxiety.

Because these differences were not of greater magnitude, the possibility that existing relationships were nonlinear was considered. As Cronbach and Snow (1969, p. 17) have stated, "Especially in studies where anxiety is a pretest variable, it is found that performance under a certain treatment is advantageous for a 'middle' group and disadvantageous at the two extremes." In accordance with their recommendation "that the investigator inspect his data for likely nonlinearities," bivariate scatter diagrams were plotted of the composite predictor-criterion relationships. Visual inspection of these scatter diagrams revealed no evidence of nonlinearity. This negative finding may well have been attributable to the particular sample of *S*s investigated. The highly anxious "tail" of the distribution could easily have been cut off through standard Navy screening processes, by the rigors of basic training, and by the further "hurdles" of selection for an electronics career field.

While it was not possible to test this hypothesis with an acceptable degree of rigor, comparisons were made between means and standard deviations of the experimental sample and normative data included in the test manuals for the Eysenck Neuroticism scale (Form B) and the two 16 PF (Form B) scales. The sample mean for the

Eysenck Neuroticism scale ($\bar{X}$ = 9.0) was found to be significantly lower (p < .01) than that of the "American College Student" norm group ($\bar{X}$ = 10.9), but none of the other means or standard deviations showed similar differences.

Other potentially meaningful results involved the unusually high criterion correlations of (1) the Set Toward Education test for the Celestial Navigation Inductive Rote course and (2) the Kuder Computational and Scientific scales for the Celestial Navigation Deductive Understanding course. While these correlations may not have been meaningful, their occurrence on a chance basis was extremely remote. For the sample size tested, Fisher's Exact Probability Test indicates that a correlation of .60 would occur by chance between one in two million and one in five million times.

There did not appear to be any logical way to account for these large correlations. In fact, were one to predict a high correlation between the Set Toward Education test and criterion performance based on what the test is presumed to measure, it would be for one of the "understanding" treatments — not for a rote learning situation. In any case, while the meaning of these high correlations was not at all obvious, it would certainly seem appropriate to include the measures which produced them in future learning style studies in hopes of gaining additional insights.

The significant negative findings of the present study related to its failure to duplicate, or even closely approximate, the results obtained in the earlier Tallmadge et al. (1968) study. The Aircraft Recognition courses used in the present study were identical to those used in the earlier study, yet none of the five statistically significant differences between inductive and deductive course predictor-criterion correlations which were found in the earlier study stood up under replication.

In an attempt to reconcile these differences, raw data and computations from the two studies were reexamined for possible errors, but none could be found. The only other factor which could be identified as possibly responsible for the differences was the fact that the Inductive Aircraft Recognition course was taught by one instructor in the earlier study while a different instructor taught the deductive course. In the present study, both Aircraft Recognition courses were taught by a single instructor.

This difference between the two studies would constitute a reasonable explanation of their different findings under some conditions. Since student-instructor interaction was kept to a minimum, however, and since the instructor in both instances was highly "programmed," this explanation did not appear satisfactory. It seemed equally unlikely, however, that 5 of the 13 correlational differences involving noncognitive predictors could have attained statistical significance on the basis of chance alone. Since at least one other investigator has found evidence of teacher by student interactions (Heil & Washburne, 1961), the former of the two explanations is perhaps the more plausible.

It was not really possible to compare the Celestial Navigation portion of the present study with the Transportation Technique portion of the previous study both because of differences between the two subject matters and because of the confounding of types of teaching with types of learning in the 1968 research. Despite these theoretical limitations, however, the fact that none of the statistically significant correlational differences of the earlier study recurred in the present study was somewhat distressing. In designing the present study, the authors considered that the Celestial Navigation and Transportation Technique topics were equivalent in all relevant respects.

Because all individual difference measures which showed significant interactive relationships in the 1968 study were included in the study reported here, it was possible, at least partially, to cross-validate the findings of the earlier study (although they did not stand up to this cross-validation). The reverse, unfortunately, was not at all feasible since the measures which produced apparently meaningful interactions here had not been included in the earlier study.

B. Relationship to Other Published Research

Other published studies have also dealt with learner anxiety levels and instructional treatment conditions seemingly similar to those investigated here. It was, however, extremely difficult to determine whether these published research findings were consistent or inconsistent with the results reported here. There were three major reasons for this difficulty. The first concerned the instructional treatments themselves which, for the most part, have been inadequately labeled and described. This problem was discussed extensively by Tallmadge and Shearer (1968), who concluded that it was frequently not possible to tell whether an instructional treatment was inductive or deductive despite its label.

A second problem related to the specific individual difference measures used. Studies purporting to deal with anxiety have used a variety of different instruments to measure it, reflecting what Cronbach and Snow (1969, p. 159) describe as "a most haphazard mixture of defining variables for the personality construct." Then too, investigators have typically used a relative criterion such as the class median to sort *S*s into high and low groups. This practice further complicates efforts to make comparisons across studies where the class medians may vary considerably.

The possible existence of undetected nonlinearities in relationships between learner personality characteristics and achievement presented another difficulty in interpreting published findings. While no evidence of nonlinearity was found in the present study, this characteristic frequently occurs where *Ss* representing the extremes of the anxiety scale are included in the treatment groups.

Studies investigating relationships between anxiety and achievement as a function of instructional treatment variables are cited by Cronbach and Snow (1969), and by Tallmadge et al. (1968). None of the cited studies, however, investigated both anxiety and inductive vs. deductive methods of instruction. It was indirect evidence which led Cronbach and Snow to the tentative conclusion that anxious *Ss* might profit more from structured than unstructured instructional treatments while the opposite might be true for less anxious *Ss*.

This conclusion would appear to be at variance with the experimental findings reported here since the three deductive experimental courses were certainly more highly structured than the three inductive courses.

If, however, one assumes that the high anxiety group in this experiment was really a "middle" group when compared against absolute standards, then the results of the present study appear more consistent with the reported literature and with common sense expectations. Because of curvilinearity of regression, both extremes of the anxiety distribution should perform better under structured conditions while the middle group should be at the optimum arousal level to both profit from the freedom and cope with the ambiguity of the less structured situation.

In summary, it must be said that there is little in the published literature which either detracts from or lends credence to the findings of the present study. While the consistency of the

obtained results across different treatment conditions increases one's confidence in the "statistical significance" of them, only future replication can be depended upon to establish the reliability of the relationships.

C. Statistical versus Practical Significance

The statistically significant findings of this study resulted from quite small mean achievement differences produced by different instructional treatments. The magnitude of these differences was clearly not sufficient to have any practical implication for Navy training programs. If the observed relationships were to stand up under field conditions, some training effectiveness gain could be expected. The cost of implementing two alternate versions of every existing course, however, would far more than offset this gain.

It must be concluded, therefore, that the results of the present study have not been of any immediate practical significance to Navy training problems. This conclusion should not, however, be taken as an indictment of the entire learning style concept. Research in this field, although prolific, has not yet become systematic. Investigators continue to be surprised by the results they obtain, and new and potentially relevant psychological parameters are serendipitously unearthed with regularity. The difficulty dimension, for example, has not yet been the subject of a single controlled experiment, although the findings of several investigators cited earlier suggest it may play a critical role in learning style research.

Similarly, the instruments used to measure potential correlates of learning styles have all been developed for different purposes. No study known to the authors of this report has even gone so far as to examine individual test items in this context.

At the present time, research in the learning style area has produced statistically significant findings. Mean achievement differences in some of these studies may even have been of sufficient size to have practical value. Still, it is not known whether these findings would stand up under replication -- much less whether they could be generalized to different subject matters and different student populations. Much more research must be done.

While the findings of some recent studies have not been particularly encouraging with respect to the eventual practical utility of individualizing instruction to match learning styles, the next generation of research projects may very well prove far more noteworthy. It can certainly be said that much has been learned about how to conduct this type of study, and the embarassingly frequent blunders detectable by hindsight should not recur.

D. A Methodological Note

A number of authors reporting research in the area of learning styles have used correlational data to test hypotheses regarding interactions between learner characteristics and treatment variables. Where differences between predictor-criterion correlations have been found as a function of treatment differences, meaningful interactions have been claimed.

Other authors, notably Cronbach and Snow (1969), have been particularly harsh in reviewing studies reporting correlational data. In discussing a study by Smith (1962), for example, they describe as "a major fault" the fact that "correlations rather than regression slopes are interpreted [pp. 90, 92]." They further describe the study as "a particularly distressing example of a painstaking and laborious experiment rendered worthless by faulty analysis [p. 92]."

In the same report, Cronbach and Snow said of the precursor to the study reported here (Tallmadge et al., 1968) that "the unavail-

ability of regression data for single aptitude variables makes detailed interpretation impossible at this time." This statement did not seem entirely credible and a decision was made to investigate the issue.

The absolute values of correlation or regression coefficients are not of primary concern in interaction studies. Rather, differences between pairs of these variables are the items of critical concern.

When testing differences between correlation coefficients, Fisher's z' transformation is used to normalize the data. When correlations are small, however, z' is very close to its corresponding r and $(z'_1 - z'_2)$ is even closer to $(r_1 - r_2)$. Since this type of research is generally dealing with both relatively small correlations and relatively small differences between correlations, $(r_1 - r_2)$ is highly correlated with the statistical significance of the difference. With $N_1 = N_2 = 53$, for example, a correlational difference of .39 is required for statistical significance ($p < .05$) when correlation coefficients are equally but oppositely different from zero, as opposed to a difference of .37 when one of the correlation coefficients is zero.

While visual inspection of differences between correlation coefficients can yield close approximations of the meaningfulness of these differences, the same is not true of differences between regression coefficients.

The formula for a regression coefficient, b, is $b = r \frac{s_y}{s_x}$ where r = the correlation coefficient, s_y = the standard deviation of the criterion score, and s_x = the standard deviation of the predictor score. Assuming a difference of, say, +.10 to -.10 between predictor-criterion correlation coefficients as a function of different instructional treatments, it is easy to see that this difference can apparently be magnified where criterion score variance (s_y^2) exceeds pre-

dictor score variance (s_x^2) and diminished when predictor score variance exceeds criterion score variance. A few examples from the present study will help to illustrate this point.

The Eysenck Lie scale yielded criterion correlations of -.09 for the Celestial Navigation Inductive Understanding treatment and +.01 for the Celestial Navigation Deductive Rote treatment. The small difference between these correlations was obviously meaningless, yet the difference between the corresponding regression coefficients (-.43 and +.05) was misleadingly large. The critical ratio for this difference was only .52. Similarly, the 16 PF More Intelligent scale produced criterion correlations of .15 and .34 respectively for the Celestial Navigation Deductive Understanding and Deductive Rote treatment groups, while the corresponding regression slopes were .97 and 1.88. Again, this seeming large difference was not statistically significant (critical ratio = .27).

An opposite kind of relationship was found for the Kuder Scientific scale which produced criterion correlations of .52 and .08 respectively for the Celestial Navigation Deductive Understanding and Aircraft Recognition Inductive treatment groups. The large difference between these correlations appeared to be meaningful and was found to be statistically significant. The small difference between the corresponding regression coefficients (.33 and .05) was not apparently noteworthy, but it, too, was found to be statistically significant (critical ratio = 2.42).

These few examples should make it abundantly clear that differences between correlation coefficients are far easier to interpret than differences between regression coefficients <u>when standard deviations are not presented</u>. Indeed, the only time when regression coefficients can yield information which is prima facie more meaningful than correlation coefficients is when the treatment has a significant effect on the variance of criterion scores and the variance

of predictor scores is identical to that of criterion scores (both groups pooled).

Since the regression coefficient, b, equals $r \frac{\sigma_y}{\sigma_x}$ and since σ_x should not differ from treatment group to treatment group (assuming random assignment of Ss), only differences in σ_y could enhance the meaningfulness of differences between regression coefficients over differences between correlation coefficients. Such differences in σ_y's could certainly result from differences in instructional treatments. It has frequently been claimed, for example, that small step programmed instruction reduces the effects of aptitude differences observed following conventional or even large step programmed instruction. While existing evidence may not be sufficient to support this hypothesis, few psychologists would question the potential of any instructional treatment to affect the variance of achievement scores as well as their mean.

As long as the possibility exists that instructional treatment conditions may affect the variance of achievement means, this kind of treatment impact should be carefully investigated and reported. Such differences are potentially more significant than differences between mean achievement scores which may be due to many extraneous factors (e.g., the deductive course writer was a "better" teacher than the inductive course writer).

As has frequently been the case, the theory that something may be — and its existence in fact — did not coincide as far as data from the present study were concerned. Correlations were computed (and are listed below) between the critical ratios of differences between z' transformations of correlation coefficients and the critical ratios of differences between corresponding regression coefficients for the following groups:

- Celestial Navigation Inductive Understanding vs. Celestial Navigation Deductive Understanding ($r = .9368$).

- Celestial Navigation Inductive Understanding vs. Celestial Navigation Inductive Rote ($r = .9320$).
- Celestial Navigation Inductive Rote vs. Celestial Navigation Deductive Rote ($r = .8998$).
- Celestial Navigation Inductive Rote vs. Aircraft Recognition Inductive ($r = .9866$).
- Celestial Navigation Inductive Rote vs. Aircraft Recognition Deductive ($r = .9763$).
- Celestial Navigation Deductive Rote vs. Aircraft Recognition Inductive ($r = .9569$).
- Celestial Navigation Deductive Rote vs. Aircraft Recognition Deductive ($r = .9227$).

These data indicate that correlational differences, which might be labeled by Cronbach and Snow as "worthless," account for at least 81%, as much as 97%, or an average of 89% of the variance of statistics which meet their criteria.

Examining data from this study's precursor yielded highly similar results. The following correlations were obtained between critical ratios of a and regression coefficient differences:

- $r = .9531$ (Transportation Technique Inductive vs. Transportation Technique Deductive).
- $r = .9766$ (Transportation Technique Inductive vs. Aircraft Recognition Inductive).
- $r = .9889$ (Aircraft Recognition Inductive vs. Aircraft Recognition Deductive).

One must conclude that correlational differences are at least highly related to the ultimate "truth" (as defined by Cronbach and Snow) of differences between regression slopes.

One further point was mentioned earlier but is deserving of further comment here. It concerns the difference between ordinal and disordinal interactions. An interaction is disordinal if the

regression line for one treatment crosses that for another within the range of experienced scores.

To the statistically sophisticated reader, it should be obvious that if mean achievement scores are identical for two treatment groups but regression slopes are different -- then any interaction must be disordinal. If the "main" effect of any experimental variable is different from zero, however, the regression lines are less likely to cross.

As was pointed out above, factors such as the respective teaching skills of different course writers may produce between-treatment main effect differences which are essentially artificial. Such main effect differences, although artificial, may change interactions which would be disordinal in their absence into ordinal interactions. For this reason, care must be exercised in interpreting reported ordinal interactions. They may be fully as meaningful as disordinal interactions.

SECTION V

CONCLUSIONS

One statistically significant and apparently meaningful interaction was found in this study. It involved learner adjustment or anxiety level and instructional methods in such a manner that the more anxious learners performed significantly better under inductive instructional conditions while the less anxious learners performed better under deductive instructional conditions (although this latter difference did not attain the $p < .05$ level of statistical significance). This interactive pattern was consistent, although not necessarily statistically significant, for three of the four adjustment-anxiety measures included in the study for two subject matters and for two types of learning (understanding and rote).

Other "statistically significant" predictor-criterion correlational differences were found when comparing different instructional treatments. Because a large number of such comparisons was made, however, and because no consistent patterns could be detected, these differences were rejected as probably not reliable.

One part of the study reported here was an almost exact replication of part of an earlier study in this series (Tallmadge et al., 1968). Not one of the statistically significant differences reported in the earlier study recurred in this research, however. The possible occurrence of a student-teacher interaction effect in the earlier study which could not have occurred in the present study (because only one teacher was involved) was listed as the most likely, although not entirely convincing, cause of this discrepancy. In any case, the failure of this study to confirm the results and conclusions of the earlier study led to a more conservative and parsimonious interpretation of present study findings.

Several quite large correlation coefficients were found relating individual difference measures to criterion performance under only one treatment condition. They could not, however, be meaningfully interpreted, and it was concluded that they should not be included in multiple predictor composites, although doing so would certainly have increased the apparent "statistical significance" of the findings.

While the results of this study are certainly less exciting than the results of the earlier study appeared to be, it is felt that they will stand up under replication and, consequently, offer realistic support for the hypothesis that individual differences in learning style exist.

SECTION VI

RECOMMENDATIONS

A. Recommendations for Present Training

It has been stated elsewhere in this report that the research findings were not of sufficient magnitude to support the cost-effectiveness of individualizing instruction on the basis of learning style differences. Statistically significant relationships between learner anxiety and teaching methods were found, however. It is not improbable that these relationships could be profited from in the usual classroom setting. Where, for example, standard expository techniques are not immediately successful in producing the desired learning for a significant number of students in a class, switching to an inductive instructional mode might represent the optimum strategy.

Even this recommendation is difficult to support if it is assumed that some cost is associated with its implementation. The research results produced to date have simply not been sufficiently clear-cut or convincing to offer a strong argument favoring any kind of change to present instructional practices.

B. Recommendations for Future Research

While the results of the present study are not particularly impressive, they are based on a conservative analysis and interpretation of the data. The study should, therefore, be considered as supporting the concept of learning styles even though the size of the observed effects was substantially less than had been expected on the basis of earlier research (Tallmadge et al., 1968). Further research should, and certainly will, be undertaken.

One negative finding of the present study, namely that neither subject matter nor type of learning appeared to be related to learner characteristics by instructional methods interactions, apparently

negates the potential utility of further investigation of these variables. While the authors would not accept this position in the absence of replication or other research specifically concerned with subject matter variables, they would concur that investigations of learner-centered variables appear more promising at the present time.

Nearly all research in the learning style area has employed individual difference measures developed for other, usually specific, applications. While those studies which tested hypotheses about aptitude variables have been well served by such instruments, studies concerned with noncognitive learner characteristics, especially personality variables, have been less fortunate. For the most part, studies of the latter type have been exploratory rather than hypothesis-oriented. They have produced statistically significant findings with at least the same frequency as aptitude-related studies but have been difficult to interpret, and especially to compare, because of possible nonlinearity of regression, differences in "class medians" between studies, and other contaminating factors.

It is the authors' belief that the most significant future developments in the learning style area will involve noncognitive learner characteristics. They also believe that specially developed measuring instruments based either on item analyses of existing tests or on specifically formulated hypotheses will be most likely to produce meaningful results.

Beswick (1969) has recently pointed out certain parallels between the learning style research reported by Tallmadge and Shearer (1969) and some of his own research on curiosity (Beswick, 1964; 1968). He has suggested that possible interactions between intrinsic motivation (curiosity) and the conditions of instruction may be attributable to "conceptual conflict" with both subject matter and instructional method acting as moderator variables.

According to Beswick's theory, Ss high on the curiosity continuum will be positively motivated by instructional treatment conditions high in "conceptual conflict" -- conditions which might be described as intellectually challenging -- and negatively motivated by low "conceptual conflict." The reverse pattern would be predicted for Ss low in curiosity. Clearly, relationships between conceptual conflict and type of learning, type of subject matter, and method of instruction would not necessarily be unidirectional, and the latter variables could interact in many ways to affect the conceptual conflict of an instructional treatment.

Beswick's hypothesis regarding the motivational effects of conceptual conflict on Ss high and low in curiosity would readily account for the results observed by Bunderson (1969) and by Burton and Goldbeck (1962) as a function of treatment difficulty differences. Further investigation might reveal other apparently conflicting results which could be explained by his conception of curiosity.

Whether or not Beswick's complex formulation of the learning style issue proves valid, it is the kind of theoretical approach which the authors believe holds promise of leading eventually to a resolution of the many still unanswered questions. They feel it is now clear that simple solutions will not be found and that, unless research is designed to test specific hypotheses, much effort will be wasted in speculation over the many seeming inconsistencies in the research literature.

REFERENCES

Armstrong, Jenny R. The relative effects of two forms of spiral curriculum organization and two modes of presentation on mathematical learning: An exploration of cancelling learner characteristic interactions. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1969.

Beswick, D. G. Notes on learning styles, instructional methods, and learning experiences. Unpublished manuscript. Australian National Univerraity, 1969.

Beswick, D. G. Correlates of individual differences in curiosity. Paper presented at the annual conference of the Australian Psychological Society, Brisbane, 1968.

Beswick, D. G. Theory and measurement of human curiosity. Unpublished doctoral dissertation, Harvard University, 1964.

Bracht, G. H. The relationship of treatment tasks, personological variables, and dependent variables to aptitude-treatment interactions. Unpublished doctoral dissertation, University of Colorado, 1969.

Bunderson, C. V. Ability by treatment interactions in designinc instruction for a hierarchical learning task. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1969.

Buros, O. K. (Ed.) *The sixth mental measurements yearbook.* Highland Park, N. J.: Gryphon Press, 1965.

Burton, B. B., & Goldbeck, R. A. *The effect of response characteristics in multiple choice alternatives on learning during programed instruction.* San Mateo, Calif.: American Institutes for Research, February 1962. AIR-C28-2/62-TR.

Capretta, P. J., Jones, R. L., Siegel, L., & Siegel, Lila C. Some noncognitive characteristics of honors program candidates. *Journal of Educational Psychology,* 1963, *54,* 268-276.

Cattell, R. B., & Eber, H. W. *Supplement of norms for Forms A and B of the Sixteen Personality Factor Questionnaire.* Champaign, Ill.: Institute for Personality and Ability Testing, 1962.

Cronbach, L. J. The logic of experiments on discovery. In L. S. Shulman and E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal.* Chicago: Rand McNally, 1966.

Cronbach, L. J., & Snow, R. E. Individual differences in learning ability as a function of instructional variables. Final Report, March 1969, Stanford University, Contract OEC 4-6-061269-1217, U. S. Office of Education.

Eysenck, H. J., & Eysenck, Sybil B. G. *Manual for the Eysenck Personality Inventory.* San Diego, Calif.: Educational and Industrial Testing Service, 1968.

Farley, F. H., & Manske, Mary E. The relationship of individual differences in the orienting response to complex learning in kindergarteners. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February 1969.

Ferney, G. A. Evaluation of a Program for Learning in Accordance with Needs. Unpublished doctoral dissertation, Washington State University, 1969.

French, J. W., Ekstrom, Ruth B., & Price, L. A. *Kit of Reference Tests for Cognitive Factors.* Princeton, N. J.: Educational Testing Service, June 1963.

Heil, L. M., & Washburne, C. Characteristics of teachers related to children's progress. *Journal of Teacher Education*, 1961, *12*, 401-406.

Klausmeier, H. J., & Quilling, Mary (Eds.) *Research and development activities in R & I units of four elementary schools of Madison, Wisconsin, 1966-1967.* Madison, Wisc.: University of Wisconsin, Wisconsin Research and Development Center for Cognitive Learning, April 1968. Technical Report No. 48.

Smith, L. M. *Programed learning in elementary school: An experimental study of relationships between mental abilities and performance.* Technical Report, USOE Title VII Project 711151.01. University of Illinois, Training Research Laboratory, 1962.

Snow, R. E., Tiffin, J., & Seibert, W. F. Individual differences and instructional film effects. *Journal of Educational Psychology*, 1965, *56*, 315-326.

Tallmadge, G. K. Relationships between training methods and learner characteristics. *Journal of Educational Psychology*, 1968, *59*, 32-36.

Tallmadge, G. K., & Shearer, J. W. *Study of training equipment and individual differences.* Orlando, Fla.: Naval Training Device Center, March 1967. Technical Report No. NAVTRADEVCEN 66-C-0043-1.

Tallmadge, G. K., Shearer, J. W., & Greenberg, Anne M. *Study of training equipment and individual differences: The effects of subject matter variables.* Orlando, Fla.: Naval Training Device Center, May 1968. Technical Report No. NAVTRADEVCEN 67-C-0114-1.

Tallmadge, G. K., & Shearer, J. W. *Study of training equipment and individual differences: The effects of subject matter variables. Supplementary Report.* Orlando, Fla.: Naval Training Device Center, August 1968. Technical Report No. NAVTRADEVCEN 67-C-0114-2.

Tallmadge, G. K., & Shearer, J. W. Relationships among learning styles, instructional methods, and the nature of learning experiences. *Journal of Educational Psychology*, 1969, *60*, 222-230.

Tanner, R. T. Expository-deductive vs. discovery-inductive programming of physical science principles. Unpublished doctoral dissertation, Stanford University, 1968.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

APPENDIX

EXCERPTS FROM THE FOUR CELESTIAL NAVIGATION COURSES

(These excerpts, taken from the four Celestial Navigation courses, all deal with finding the distance from the observer's true position to the Geographical Position of a star.)

CELESTIAL NAVIGATION INDUCTIVE UNDERSTANDING COURSE

When I'm standing at the North Pole, at what altitude will Polaris appear above the horizon? Right,* it is 90 degrees. And what is the latitude of the North Pole? Right, it is also 90 degrees. When I'm standing at the North Pole, how many degrees am I away from the point on earth directly beneath Polaris? Right, I am zero degrees away. When I am standing at the equator, how many degrees is Polaris above the horizon? Right, it is right on the horizon or zero degrees above it. What is the latitude of the equator? Right, it is zero degrees. Now, — how many degrees away from the point on earth directly beneath Polaris am I when I am standing at the equator? Well, if I'm zero degrees away when I'm standing at the North Pole (and Polaris is 90 degrees above the horizon), then when I'm standing at the equator (and the altitude of Polaris is zero degrees and my latitude is zero degrees), how far away am I from the North Pole? Isn't that the same distance I am away from the point directly beneath Polaris? Right, when I am standing at the equator, the altitude of Polaris above the horizon is zero degrees, my latitude is zero degrees, and my distance from the point on earth directly beneath Polaris is 90 degrees. There's a relationship here between the measured altitude of Polaris above the horizon and my distance from the point directly beneath Polaris, the North Pole. Forget about latitude for a minute, who can tell me what the relationship is? Right, my distance from the North Pole, or the spot directly beneath Polaris is 90 degrees minus the measured altitude of Polaris above the horizon.

---

* This inductive course was designed to elicit and react to student responses. The word, "Right," is used here under the assumption that at least one student answered the preceding question correctly (and puzzled expressions did not appear on too many faces). The entire monologue from this point on would have been different if the question had not been correctly answered.

CELESTIAL NAVIGATION DEDUCTIVE UNDERSTANDING COURSE

When I am standing at the North Pole, Polaris is 90 degrees above the horizon. This means that when Polaris is 90 degrees above the horizon, I am zero degrees away from the spot directly underneath Polaris. If I am standing at the Equator, Polaris is zero degrees above the Equator, but I am 90 degrees away from the spot directly beneath Polaris. Remember, the latitude of the North Pole is 90 degrees north. The latitude of the Equator is zero degrees. Any place I may be in the Northern Hemisphere, I can find how many degrees I am away from the North Pole by subtracting my latitude from 90 degrees. If I am at the North Pole, I subtract my latitude, 90 degrees, from 90 degrees and get zero degrees. This means that I am zero degrees away from the spot on earth directly beneath Polaris. If I am at the Equator, I can subtract my latitude, zero degrees, from 90 degrees and find that I am 90 degrees away from the spot on earth directly beneath Polaris. If my latitude is 40 degrees, I can subtract that from 90 degrees and find that I am 50 degrees away from the North Pole, the spot on earth directly beneath Polaris. Now remember that latitude is equal to the measured height of Polaris from the horizon, so that to find out how far I am away from the point on earth directly beneath Polaris, I measure the height of Polaris above the horizon and subtract that from 90 degrees.

Another navigation problem which we will cover today has to do with finding the true distance from a specific location to the Geographical Position of a star, when you know the altitude of the star. Suppose that the altitude of Vega, measured from your location, is 77 degrees. How far are you from the GP of Vega?

(Write the following example on the chalkboard:)

Altitude of Vega = 77°

$$\begin{array}{r} 90^\circ \\ -\ \underline{77^\circ} \\ 13^\circ \end{array} \qquad \begin{array}{r} 13 \\ \times\ \underline{60} \\ 780 \end{array} \text{ nautical miles}$$

So you are locat[illegible] 80 nautical miles from the GP of Vega. Let's try another exam[illegible]e. Suppose the altitude of Pollux, measured from your location, is 84 degrees, 20 minutes. How far are you from the GP of Pollux?

(Write the following on the chalkboard:)

Altitude of Pollux = 84° 20'

$$\begin{array}{r} 90^\circ\ 00' \\ -\ \underline{84^\circ\ 20'} \end{array}$$

Since 1° = 60 minutes we can write our problem like this:

$$\begin{array}{r} 89^\circ\ 60' \\ -\ \underline{84^\circ\ 20'} \\ 5^\circ\ 40' \end{array} \qquad \begin{array}{r} 5 \\ \times\ \underline{60} \\ 300 \end{array} \qquad \begin{array}{r} 300 \\ +\ \underline{40} \\ 340 \end{array} \text{ nautical miles}$$

So you are 340 nautical miles from the GP of Pollux. Let's look at one more example. Suppose you measure the altitude of Altair and find it to be 60 degrees, 14 minutes. We can write our problem like this:

$$\begin{array}{r} 90^\circ\ 00' \\ -\ \underline{68^\circ\ 14'} \end{array}$$

Since there are 60' in a degree, we can write our problem like

CELESTIAL NAVIGATION DEDUCTIVE ROTE COURSE

Another navigation problem which we will cover today has to do with finding the true distance from a specific location to the Geographical Position of a star, given the altitude of the star. To find that distance you first subtract the altitude of the star above the horizon (which you would ordinarily measure with a sextant) from 90 degrees, and then convert the obtained difference into nautical miles. Suppose that the altitude of Vega, measured from your location, is 77 degrees. How far are you from the GP of Vega? First, you would subtract 77 degrees from 90 degrees to get 13 degrees. The next step is to convert 13 degrees to nautical miles. It just so happens that 1 degree equals 60 nautical miles, so you would multiply the number of degrees by 60. So, 13 x 60 = 780 nautical miles. It is 780 nautical miles from the spot where you measured the altitude of Vega to the GP of Vega (which is the point directly beneath Vega). Let's try another example. Suppose you measure the altitude of Altair and find it to be 60 degrees, 14 minutes. Let's look at how we do that. We can write our problem like this:

$$\begin{array}{r} 90^\circ\ 00' \\ -\ \underline{68^\circ\ 14'} \end{array}$$

Since there are 60 minutes in a degree, we can borrow a degree from 90 degrees and convert it to 60 minutes. Now our problem looks like this:

$$\begin{array}{r} 89^\circ\ 60' \\ -\ \underline{68^\circ\ 14'} \\ 21^\circ\ 46' \end{array}$$

and we can proceed with our subtraction to obtain 21 degrees, 46 minutes. Since 1 degree equals 60 nautical miles and 1 minute equals 1 nautical mile, to convert 21 degrees, 46 minutes to nautical miles, we simply multiply the number of degrees by 60 and then add to that the number of minutes. 21 x 60 = 1260 and to that we add 46 to obtain 1306 nautical miles, which is the distance from our location to the GP of Altair. You subtract the altitude of the

UNCLASSIFIED

DOCUMENT CONTROL DATA - R & D

American Institutes for Research Palo Alto, California	Unclassified

STUDY OF TRAINING EQUIPMENT AND INDIVIDUAL DIFFERENCES: Research on Interactive Relationships Among Learner Characteristics, Types of Learning, Instructional Methods, and Subject Matter Variables

Technical Report 28 June 1968-28 February 1970

G. Kasten Tallmadge
James W. Shearer

June 1970	vi, 67	27

N61339-68-C-0271 7690-4	Technical Report NAVTRADEVCEN 68-C-0271-1

This report will carry Distribution Statement No. 2.

None	Human Factors Laboratory Naval Training Device Center Orlando, Florida 32813

This is the fourth in a series of research projects aimed at determining whether learning might be enhanced by employing instructional methods which differ in design and use as a function of learner characteristics. Based on inference drawn from studies in this series and other research literature, a model was developed enabling the simultaneous examination of the effects of learner characteristics, types of learning, instructional methods, and subject matter variables on achievement. Each of six experimental courses was administered to ~~between 57 and 60~~ Navy enlisted men who were previously tested with instruments which yielded a total of 35 measures of aptitude, interest, and personality characteristics. Correlation and regression analyses revealed no consistent and meaningful interactive relationships existing between learner characteristics and types of learning or types of subject matter. These analyses did, however, reveal an apparently consistent and meaningful interaction between learner anxiety level and method of instruction (inductive vs deductive). While the magnitude of the observed relationship was not sufficient to give promise of immediate practical application, it was concluded that the research supported the existence of individual differences in learning style.

DD FORM 1473 (PAGE 1)

UNCLASSIFIED

Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Individual differences						
Learning style						
Deductive instruction						
Inductive instruction						
Instructional methods						
Training equipment						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

UNCLASSIFIED
Security Classification